

Analyseur de log – Analyse

Vocabulaire utilisé.....	2
Introduction	2
Brève présentation	3
Description des données récupérées	4
Données musculaires du visage (Expressiv suite).....	4
Données émotionnelles (Emotiv suite).....	5
Données cognitives (Cognitiv suite)	5
Données gyroscopes	5
Analyses des données.....	6
Valeurs émotionnelles.....	6
Données expressives, cognitives et gyroscope	9
Eléments de comparaison.....	9

Vocabulaire utilisé

- EPOC : Il s'agit d'un casque d'interfaçage cerveau-ordinateur conçu par l'entreprise Emotiv. Il se compose de 16 électrodes humides et 2 gyroscopes. En plus des signaux cérébraux, Emotiv propose la mesure de certains états émotifs ainsi que les expressions du visage. [Site Internet](#)
- chdh : « chdh est une performance proposant une symbiose entre le son et l'image, où l'un et l'autre se complètent jusqu'à devenir indissociables. » [Site internet](#)
- Modalités : On désigne par modalités les différentes actions présentées et détectées par le casque EPOC d'Emotiv. Par exemple : le clin d'œil, la frustration, le sourire...
- Joysticks : On désigne par joysticks les différentes modalités lorsque celle-ci sont utilisées pour piloter quelque chose.

Introduction

Dans le cadre du [NIFFE](#), la Haute Ecole Arc a décidée de présenter une façon de mettre en image et en son l'activité cérébrale de tout un chacun. Ce démonstrateur fait l'objet de deux projets de semestres consécutifs dont voici résumés les objectifs :

1. Etudes des différentes modalités proposées par le casque EPOC de chez Emotiv afin de déterminer lesquelles fonctionnent le mieux et le plus rapidement. Une modalité est par exemple le clin d'œil gauche, mais il y a aussi serrer les dents, un haut niveau de frustration...
2. Interfaçage du casque EPOC d'Emotiv avec l'application [chdh](#)

Le premier projet propose de soumettre à différents testeurs divers stimuli et d'enregistrer les données mesurées par le casque EPOC. Ces données sont ensuite analysées afin de mettre en évidence les modalités les plus efficaces chez le plus grand nombre d'individus.

La présente analyse concerne l'application chargée d'analyser les données récoltées. Il s'agit en outre de déterminer les types de données, comment les traiter pour obtenir :

- Les modalités les plus répondantes sur un maximum de personne

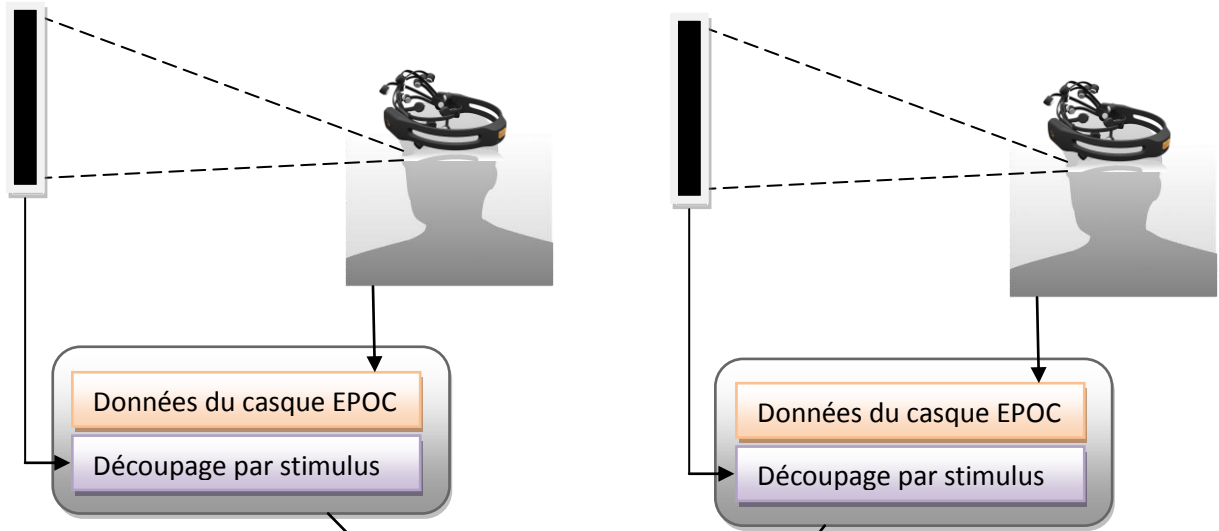
Mais aussi :

- Indépendantes entre elles, c'est-à-dire aucune corrélation statistique entre l'activation des différentes modalités
- Pour ces modalités, l'utilisation de seuil est-il nécessaire ? Si oui, quelle valeur pour le plus grand nombre de personne

Brève présentation

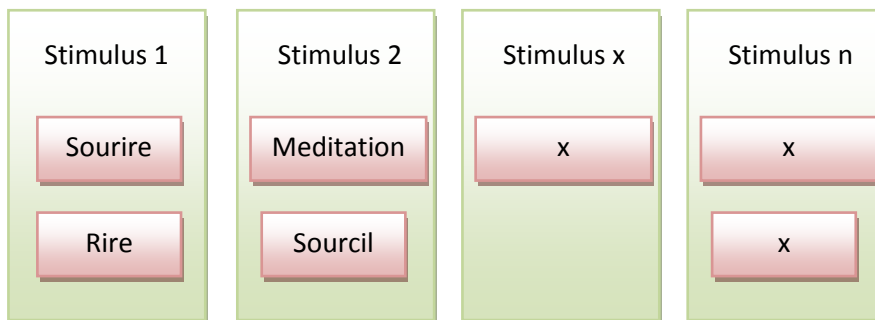
Présentation de n stimuli

Présentation de n stimuli



Le principe est donc de présenter des stimuli visuels et d'enregistrer les données captées par le casque. Il s'agit ensuite d'analyser ces données afin d'identifier la ou les modalités les plus réactives.

Analyse

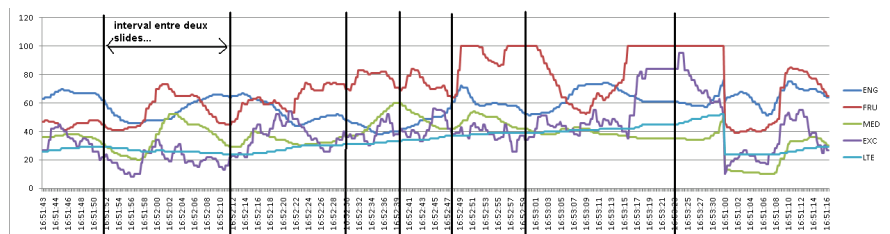


Au final, pour chaque stimulus, nous obtenons la/les modalités les plus réactives et il faut identifier celles qui ont le moins de corrélation entre elles.

Description des données récupérées

Les données récupérées proviennent de deux applications différentes. Les données du casque EPOC sont loguées à l'aide de l'application « NeuroVault » disponible sur le site internet d'Emotiv. Les données enregistrées, plus précisément, celles utilisées dans le cadre de cette analyse sont regroupées en quatre catégories : (1) Données musculaires, (2) données émotionnelles, (3) données cognitives et (4) données des gyroscopes. Outre les modalités de la première catégorie -qui sont rapides car musculaires – ainsi que des données gyroscopiques, les autres modalités sont plus « subjectives » dans le sens où nous ne savons pas comment ces modalités sont déterminées. Cependant, les données émotionnelles et cognitives sont le reflet de l'activité cérébrale, réel objet du travail et du casque EPOC.

La deuxième application est une extension développée pour Powerpoint qui enregistre l'heure à laquelle chaque page du diaporama change. Chaque page propose une action ou activité différente, les données en contiennent donc l'heure de début et de fin.



Ces deux sources de données permettent d'analyser les données de chaque modalité pendant le temps écoulé sur chaque page du diaporama (pour chaque stimulus).

Données musculaires du visage (Expressiv suite)

Les modalités mesurées et loguées peuvent être regroupées en deux catégories. Pour ces deux catégories, lorsque le mouvement n'est pas détecté, la valeur loguée est à zéro. Une valeur différente de zéro indique qu'il y a détection ainsi que son intensité.

Pour résumer :

valeur <> 0 = détection / valeur = intensité (puissance)

A. Catégorie dans l'intervalle [0 ; 100]

SMILE : sourire, LAUGH : rire, CLENCH : serer les dents (grogner), SMIRKL, SMIRKR : sourire de côté gauche respectivement droite, WINKL, WINKR : Clin d'œil gauche respectivement droite

B. Catégorie dans l'intervalle [-100 ;100]

LOOKH : Mouvement des yeux horizontal (valeur < 0 = gauche, valeur > 0 = droite)
LOOKV : Mouvement des yeux vertical (valeur < 0 = down, valeur > 0 = up)
BROW : Mouvement des sourcils (valeur < 0 = down, valeur > 0 = up)

Données émotionnelles (Emotiv suite)

Les modalités mesurées et loguées sont : **ENG** (engagement), **FRU** (frustration), **MED** (meditation), **EXC** (excitement) et **LTE** (long term excitement)

Les valeurs loguées sont comprises dans l'intervalle [0 ; 100]. Contrairement aux données musculaires, ces valeurs sont globalement supérieur à 0 et oscillent dans une certaine plage de valeur.

Détecter une activité de l'une de ces modalités revient donc à déterminer si son intensité à changée, vers le haut ou vers le bas par rapport à la séquence précédente. Il est prévu de découper le log en séquences déterminées par le moment de changement de diaporama du test.

Données cognitives (Cognitiv suite)

Pour que des données cognitives soient enregistrées, cela nécessite un temps de calibrage. De plus, au maximum 4 des modalités peuvent être calibrés par utilisateur. Si aucun calibrage n'a été fait, les valeurs sont égales à 0.

Le calibrage consiste à tenter d'effectuer une action par la pensée et le programme tentera d'identifier ces signaux une prochaine fois.

Les modalités mesurées et loguées sont : PUSH (pousser), PULL (tirer), LIFT (soulever), DROP (déposer), RIGHT (droite), LEFT (gauche), ROTR (rotation droite), ROTL (rotation gauche), ROTCW (rotation horaire), ROTCCW (rotation antihoraire), ROTFWD (rotation avant), ROTREV (rotation arrière) et VANISH (disparaître).

Les valeurs loguées sont comprises dans l'intervalle [0 ; 100].

Une valeur différente de zéro indique qu'il y a détection ainsi que son intensité.

Pour résumer :

valeur > 0 = détection / valeur = intensité (puissance)

Données gyroscopes

Les modalités mesurées et loguées sont :

GYROX : valeur < 0 = déplacement gauche, valeur > 0 = déplacement droite

GYROY : valeur < 0 = déplacement bas, valeur > 0 = déplacement haut

Les valeurs loguées sont comprises dans l'intervalle [-3000 ; 3000]. Cet intervalle n'a pas pu être vérifié dans la documentation ni ailleurs pour l'instant.

Une valeur différente de zéro indique qu'il y a détection ainsi que son intensité.

Pour résumer :

valeur <> 0 = détection / valeur = intensité (puissance)

Analyses des données

Voici deux tableaux récapitulatifs des données, de leur plage de valeur et de la condition pour laquelle elles sont nulles.

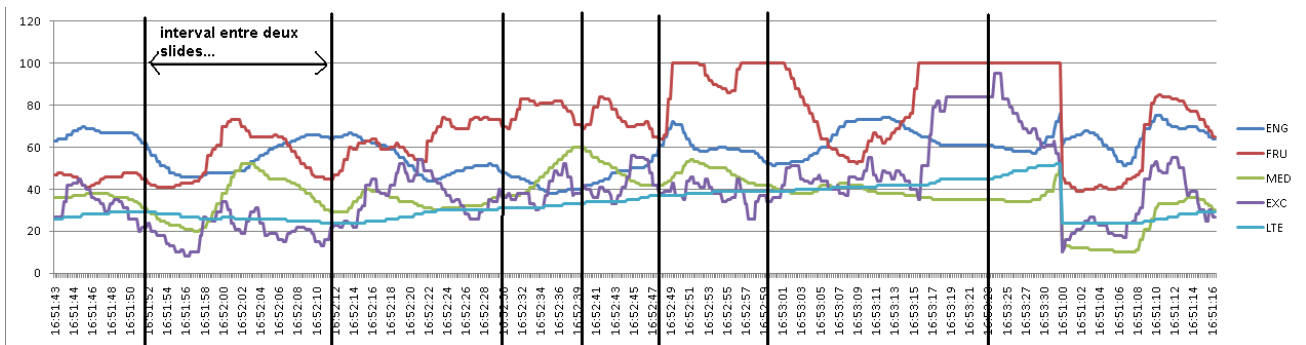
	GYROX	GYROY	ENG	FRU	MED	EXC	LTE	SMILE	LAUGH	CLENCH	SMIRKL	SMIRKR	WINKL	WINKR	LOOKH	LOOKV	BROW
min	-3000	-3000	0	0	0	0	0	0	0	0	0	0	0	0	-100	-100	-100
max	3000	3000	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
NO-DETECT	0	0	---	---	---	---	---	0	0	0	0	0	0	0	0	0	0
positive	right	up	power	power	power	power	power	power	power	power	power	power	power	power	right	up	up
negative	left	down	---	---	---	---	---	---	---	---	---	---	---	---	left	down	down

	PUSH	PULL	LIFT	DROP	RIGHT	LEFT	ROTR	ROTL	ROTCW	ROTCCW	ROTFWD	ROTREV	VANISH
min	0	0	0	0	0	0	0	0	0	0	0	0	0
max	100	100	100	100	100	100	100	100	100	100	100	100	100
NO-DETECT	0	0	0	0	0	0	0	0	0	0	0	0	0
positive	power	power	power	power	power	power	power	power	power	power	power	power	power
negative	---	---	---	---	---	---	---	---	---	---	---	---	---

Valeurs émotionnelles

Contrairement aux autres données, les valeurs émotionnelles ne sont pas à zéro lorsqu'il ne se passe rien. Logique puisqu'il est rare que notre cerveau ne fasse rien et nos émotions sont difficilement contrôlables et fluctuent sans cesse. Cependant, une personne entraînée devrait pouvoir faire chuter à zéro une des émotions mesurées. Ces affirmations reposent sur la détection proposée par Emotiv. Elles supposent les modalités ENG, FRU, MED, EXC et LTE reflètent bien nos émotions.

Dans le but de détecter si un stimulus a de l'effet sur l'une des ces modalités, décrivons tout d'abord ce que l'on a.



Les données sont découpées en segments. Ces segments sont créés sur la base du logueur de la présentation (stimuli). A chaque transition d'activité (de page), le temps est enregistré. Nous possédons ainsi les données pour chaque stimulus.

Remarque : Entre le moment où le stimulus est présenté et le moment où le signal commence à changer, il peut se passer jusqu'à 10 secondes. Il faut prévoir suffisamment de temps pour chaque stimulus. Il se peut que dans le début d'un segment, un signal soit en phase changeante suite au stimulus du segment précédent.

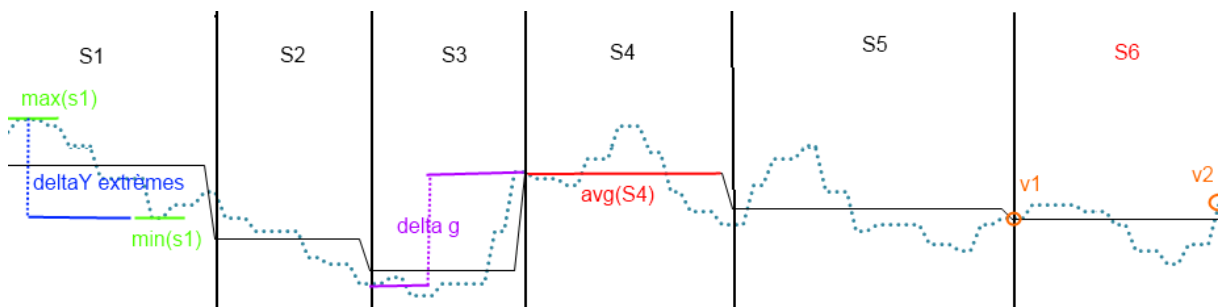
Le but est de détecter une **activité significative** dans un segment tout en sachant que dans le cas des modalités émotionnelles, il n'y a pas de **référence de « non activité »**. Il faut donc dans un premier temps créer une référence sur laquelle se baser pour déterminer si le signal mesuré est le reflet d'une **activité intéressante**.

Qu'est-ce qu'une activité intéressante ou significative ?

Lorsqu'un sujet est soumis à un stimulus externe, nous considérons comme intéressantes ou significatives, les données récoltées pour lesquelles il est possible de mesurer l'effet du stimulus. Cela se traduit par des variations de valeurs conséquentes, des minima et maxima, des moyennes élevées ou basses...

Quelle référence choisir ?

Les données du segment précédent peuvent jouer le rôle de référence. En effet, si un stimulus obtient des effets sur le segment précédent, cet effet peut soit perdurer, augmenter ou diminuer dans le segment courant.



Sur ce graphique, la courbe bleue représente les valeurs recueillies par le casque pour une modalité (exemple : ENG). Les droites horizontales représentent les moyennes de chaque segment. Les droites verticales représentent les différents segments.

Quelles valeurs pouvons-nous calculer ?

$avg(S)$ = moyenne globale

$v1$ = valeur en entrée du segment

$v2$ = valeur en sortie du segment

$avg(S_i)$ = moyenne d'un segment

$min(S_i)$ = minima d'un segment

$max(S_i)$ = maxima d'un segment

$\Delta y_{ext}(S_i) = maxY(S_i) - minY(S_i)$ = écart (vertical) entre le minima et le maxima

$\Delta y_g(S_i) = v2Y(S_i) - v1Y(S_i)$ = écart (vertical) entre l'entrée et la sortie

$f'(S_i)$ = dérivée première, pente de la courbe

$f''(S_i)$ = dérivée seconde, courbure

$V(S_i) = 1/n * \sum_{i=1}^n (x-avg(S_i))^2$ = variance du segment

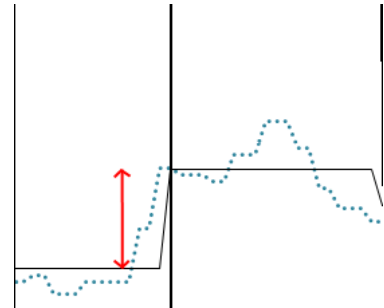
Voici une proposition de critères qui peuvent entrer en ligne de compte pour déterminer les segments les plus réactifs aux stimuli.

A. L'écart entre la moyenne de deux segments consécutifs

Plus l'écart est grand, plus leurs différences est significative de changement.

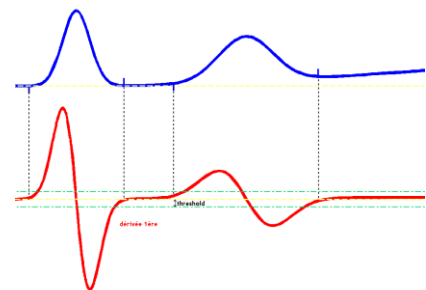
Mais attention, deux segments consécutifs peuvent présenter une moyenne proche et être très différents. Ce facteur ne doit pas être seul dans le calcul.

$$\Delta \text{avg}(S_i) = |\text{avg}(S_i) - \text{avg}(S_{i-1})|$$



B. La dérivée première

Calculer la dérivée première revient à calculer la pente de chaque point de la courbe. Ceci nous donne une indication quant à l'agitation de la courbe. Comme dans l'illustration, il s'agit d'introduire un seuil (en vert) mais également un filtre en option afin d'éliminer le bruit éventuel (moyenne des n points environnants).



$$\text{si } f'(S_i) \text{ dépasse le seuil, alors on prend la valeur } \sum_{i=1}^n |f'(x)|$$

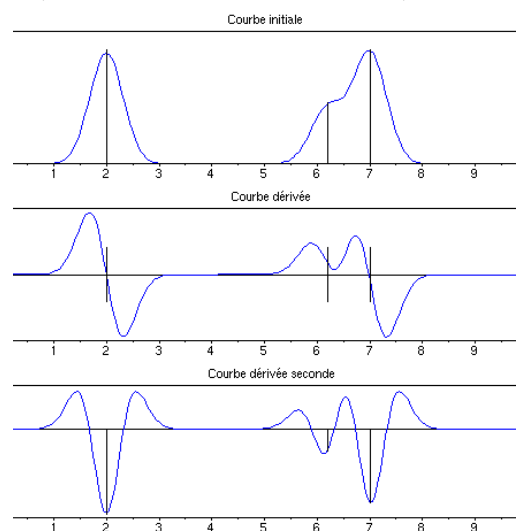
C. La dérivée seconde

Si elle est positive sur un intervalle, la pente augmente, la courbure est vers le bas, la fonction est dite « convexe » sur cet intervalle.

Si elle est négative sur un intervalle, la pente diminue, la courbure est vers le haut, la fonction est dite « concave » sur cet intervalle.

Si elle est nulle, la courbe est localement rectiligne.

Si la dérivée seconde s'annule et change de signe, on a un point d'inflexion, la courbure de la courbe s'inverse.

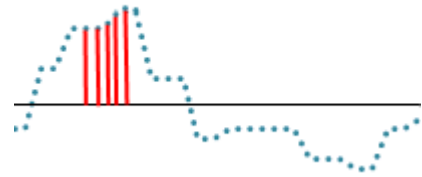


On prend en compte,

$$\sum_{i=1}^n |f''(x)| \text{ si la valeur dépasse le seuil fixé}$$

D. La variance du segment

La variance est la somme des carrés des écarts entre chaque point et la moyenne du segment. Une variance faible signifie peu d'activité sur la courbe tandis qu'une grande variance implique de grands écarts.



$$V(S_i) = 1/n * \sum_{i=1}^n (x - \text{avg}(S_i))^2$$

E. La formule

Pour terminer, l'indice d'activité du segment :

$$IA(S_i) = \Delta \text{avg}(S_i) + \sum_{i=1}^n |f'(x)| + \sum_{i=1}^n |f''(x)| + V(S_i)$$

Avec l'indication concernant la tendance du segment (signe de la somme des dérivées premières, signe de la somme des dérivées secondes).

Pour assurer une valeur entre 0 et 100, l'indice de chaque segment est calculé. L'indice est ensuite normalisé en fonction du segment au plus haut indice.

Seuils paramétrables :

- Seuil sur la valeur $\Delta \text{avg}(S_i)$. La valeur de $\Delta \text{avg}(S_i)$ est prise en compte si $>$ seuil.
- Si $f'(x)$ dépasse le seuil sur un point, la valeur $\sum_{i=1}^n |f'(x)|$ est prise en compte
- Seuil sur la valeur $\sum_{i=1}^n |f''(x)|$. La valeur $\sum_{i=1}^n |f''(x)|$ est prise en compte si $>$ seuil.
- Seuil sur $V(S_i)$, la variance est prise en compte si elle dépasse le seuil.

NB : Cette formule à été mise sur pied par une équipe non spécialisée dans l'analyse mathématique de données et établie dans l'urgence. Cette formule ne peut en aucun cas être considérée comme définitive. Plusieurs personnes la trouvent peu convaincante mais n'ont hélas pas proposé d'alternative.

Données expressives, cognitives et gyroscope

Ces données ont toutes en commun le fait qu'une valeur différente de zéro implique que la modalité a été détectée. Contrairement aux émotions, le travail de détection est déjà fait.

Il suffit de mettre en évidence les modalités qui ont la plus grande puissance. Pour ce qui est du gyroscope, cette donnée peut être échelonnée entre 0 et 100.

Eléments de comparaison

Toutes les valeurs servant à la comparaison entre les différentes modalités sont- en valeur absolue- comprises entre 0 et 100. Certaines valeurs sont exprimées dans l'intervalle [-100 ; 0] ce qui en valeur absolue revient au même que l'intervalle [0; 100].

En calculant l'indice d'activité de chaque modalité pour chaque segment, nous avons pour chaque segment : IA(Modalité1), IA(Modalité2)...IA(ModalitéN).

	Segment 1	Segment 2	...	Segment n
	Stimulus 1	Stimulus 2	...	Stimulus n
Modalité 1	88	0	...	IA
Modalité 2	20	54	...	IA
Modalité 3	73	1	...	IA
...
Modalité n	IA	IA	...	IA

L'étape suivante consiste à comparer les IA d'un segment afin d'en sortir le plus grand IA significativement supérieur au prochain plus grand. Cette différence pourrait être représentée par un seuil paramétrable. Si aucun ne peut être écarté, nous pouvons prendre les n plus grands. Il est intéressant de constater si un groupe précis d'IA s'active systématiquement ensemble dans certains cas.

Nous obtenons donc, par segment (stimulus) : 0 à n modalité(s) avec un bon IA

Stimulus 1	modalité 1, modalité 3		
Stimulus 2	modalité 2		

La modalité souhaitée est-elle présente ? Si ce n'est pas le cas, mémoriser cela afin d'afficher cela visuellement. Si aucune modalité n'est souhaitée, par exemple dans le cas du casse tête pour lequel nous cherchons à mesurer les modalités actives, nous pourrions vouloir voir quelle modalité a le plus haut IA. Dans ce cas, mettre de côté tous les IA retenus jusque là.

Il faut maintenant s'assurer que la modalité retenue n'apparaît pas dans un autre segment. Ceci parce que nous souhaitons une modalité qui ait peu ou pas de corrélation avec les autres stimuli.

Comme nous recherchons les modalités les plus indépendantes, que la modalité retenue soit attendu (clin d'œil) ou qu'elle soit simplement mise en évidence (méditation pendant casse tête), sa présence à plus d'un endroit doit être signalée. Mémoriser ces informations et les afficher.

Pour terminer, comparer les modalités retenues entre tous les cobayes et mettre en évidence les plus récurrents !

Stimulus 1	modalité 1, modalité 3		
Stimulus 2	modalité 2		
Stimulus 1	modalité 1, modalité 2		
Stimulus 2	modalité 4		
Stimulus 1	modalité 1, modalité 3		
Stimulus 2	modalité 2		

Une question n'est pas traitée dans cette analyse. Que se passe-t-il si on soumet le même sujet à la même suite de stimuli plusieurs fois ?

Comme nous souhaitons connaître les « meilleures » modalités, c'est-à-dire celles qui fonctionnent le mieux chez un maximum de personnes, l'analyse actuelle suffit. Le temps manque pour répondre à la question susmentionnée et elle demeure ouverte.